Social Media and Stock Price Reaction to Data Breach Announcements: Evidence from US Listed Companies

Pierangelo Rosati[*†‡], Peter Deeney[†], Mark Cummins[†],
Lisa Van der Werff[†], Theo Lynn[†‡]

[†]*DCU Business School*
[‡]*Irish Centre for Cloud Computing and Commerce (IC4)*

ABSTRACT

Data breaches are not only on the increase but firms struggle to detect, defend and respond to such breaches. A data breach opens a period of crisis for the affected firm, generates complex information, and requires providing information to a variety of stakeholders in a timely and proper manner. This article reports one of the first studies on the impact of social media exposure by affected firms on stock price reaction to a data breach announcement. Using an event study methodology on a sample of 87 data breaches from 73 US publicly-traded firms from 2011 to 2014, we find that use of social media exposure at the time of a data breach exacerbates the negative stock price to the announcement. Interestingly, we find that this negative association is contingent on traditional media visibility; the effect is positive for low-visibility companies. Based on our results, we posit that there is a need for a contingency model for social media communication during firm crises and such a model should be based at least on firm size, visibility and the type of crisis.

Keywords: data breach; social media; stock market; corporate disclosure; firm visibility; event-study methodology.

Forthcoming in Research in International Business and Finance

[*] Corresponding author: Pierangelo Rosati, DCU Business School, Dublin City University, Glasnevin, Dublin 9, Ireland.
Email: pierangelo.rosati@dcu.ie

Acknowledgment: The research work described in this paper was supported by the Irish Centre for Cloud Computing and Commerce, an Irish National Technology Centre funded by Enterprise Ireland and the Irish Industrial Development Authority.

1. Introduction

In the current business environment, firms rely heavily on information systems and data analytics to build their competitive advantage (LaValle, Lesser, Shockley, Hopkins, and Kruschwitz 2011; Chen, Chiang, and Storey 2012). The amount of data organizations collect, store and process has grown exponentially in the last few years (LaValle et al. 2011; Liu and Ye 2016). This data usually contains valuable and sensitive information about customers, business partners, and about the organization itself. Therefore, a data loss or the involuntary disclosure of such information may generate significant damage for the affected organization. Given the rapid growth of the number the data breaches over the last few years (Ponemon Institute 2018), cyber security for firms has become a real concern for managers, investors and regulators (Ponemon Institute 2016; Verizon 2017).

In this paper, we investigate the effect of social media exposure by affected firms on stock price reaction to a data breach announcement in order to understand whether such an alternative communication channel mitigates or exacerbates the cost of a data breach. A data breach is defined as an incident that involves unauthorized access to sensitive, protected, or confidential data[1] resulting in the compromise or potential compromise of either confidentiality, integrity, or availability of an information asset (Gordon, Loeb, and Zhou 2011). The number of information systems breaches is growing every year and the increasing popularity of cloud computing, mobile devices and big data exacerbate this issue (Romanosky, Hoffman, and Acquisti 2014; Abbasi, Saker abd Chiang 2016). As such, firms are investing more and more in ways to protect their information systems from cyber-attacks (Srinidhi, Yan, and Tayi 2015). According to the Privacy Rights Clearinghouse, 543 million records were lost between January 2005

[1] Health information, personal identifiable information, trade secrets or intellectual property, and/or personal financial data are typical examples of sensitive and confidential information (Sen and Borle 2015).

and January 2012 as a consequence of 2,800 data breaches (Risius and Beck 2015). Data breaches impose significant costs on the affected companies both in the short and long term. Short-term costs are due to investigation and remediation activities, legal advisory, fines, and lost transactions (Aral, Dellarocas, and Godes 2013). A prominent example of the short-term cost of a data breach is ChoicePoint. In early 2006, the Federal Trade Commission (FTC) imposed a $10 million fine against ChoicePoint as a consequence of a massive data breach that involved 160,000 records; the company also agreed to pay another $5 million to compensate affected individuals (Federal Trade Commission – FTC 2009). Long-term costs are related to loss of present and future revenues as well as the deterioration of customers' and partners' trust (Charette, Adams, and White 1997; Cavusoglu, Mishra, and Raghunathan 2004; Aral et al. 2013; Dennis, Wixom, and Tegarden 2015). These long-term costs usually represent most of the overall cost of a data breach (Goel and Shawky 2009; Gatzlaff and McCullough 2010). However, they are extremely hard to quantify. For this reason, empirical researchers adopt stock price reaction (i.e. stock return) as a proxy (Goel and Shawky 2009; Gatzlaff and McCullough 2010) and have shown that a breach may cause a loss in firm value of up to 5.5 percent (Campbell, Gordon, Loeb, and Zhou 2003).

Given its unexpected nature and the potential harm that a data breach may generate, it clearly fits the definition of a company crisis proposed by Schultz, Utz, and Göritz (2011)[2]. In this context, timely communication can limit potential harm (Lee, Hutton, and Shu 2015) and this is one of main concerns for regulators in the case of a data breach. However, investigating a data breach is complex and may require a significant amount of time and deep technical capabilities (Casey 2006). As a result, details about the incident may not become apparent or public for some time resulting in uncertainty

[2] According to Schultz et al. (2011). a firm's crisis can be defined as 'a specific, unexpected and non-routine event or series of events that create high levels of uncertainty and threaten, or are perceived to threaten, an organization's high priority goals'.

which may adversely affect the market reaction (Kalev, Liu, Pham, and Jarnecic 2004). during this period. Similarly, the explanation surrounding the data breach announcement may be complex therefore how this information is communicated is particularly important (Coombs 2007; Utz, Schultz, and Glocka 2013).

Firms have usually relied on traditional media (i.e. the press) to disseminate information (Blankerspoon, Miller, and White 2014). However, traditional media may not always be the most useful communication channel in the context of a crisis as it tends to focus on highly visible firms since they attract larger readership (Miller 2006; Barber and Odean 2008). As a result, low visibility firms, which represent the largest part of the market, struggle in reaching their stakeholders through traditional media, and this may be particularly detrimental in the context of company crisis. The emergence of social media has signaled a step change in recent years offering firms an alternative communication channel through which they can disseminate information more effectively, to a wider audience, and at a relatively low cost.

Social media is a set of internet-based tools and applications that allow users to create (consume) content that can be consumed (created) by others and which enables and facilitates connections (Hoffman and Novak 2012). Social media is now widely adopted by firms for corporate communication and has been recognized as an official communication channel (Securities and Exchange Commission – SEC 2013). Studies suggest that (i) social media represents nowadays the main source of information for an increasing number of people (Jansen, Zhang, Sobel, and Chowdury 2009; Coombs 2014), (ii) that information disseminated through such a communication channel has impact on stock returns and information asymmetry (Blankerspoon et al. 2014; Jung, Naughton, Tahoun, and Wang 2015), and (iii) that it is an effective communication channel during company crises (Lee et al. 2015; Jahng and Hong 2017).

Among different social media platforms, many studies focus on Twitter since it is the most commonly adopted for social investor communication and company event disclosure (Blankerspoon et al. 2014; Jung et al. 2015). Furthermore, Twitter has the peculiarity of being a largely open network and it also has the unique feature of 'retweeting', which makes it a powerful mechanism for information sharing (Kietzmann, Hermkens, McCarthy and Silvestre 2011; Stieglitz and Dang-Xuan 2013).

In this study, we specifically focus on Twitter since anecdotal evidence reveals that firms try to exploit its characteristics to disseminate information about data breaches. For example, on September 3rd 2014, following a breach to their payment card system, The Home Depot (@HomeDepot) tweeted: "*To keep customers updated, we've posted a message about news reports of a possible payment data breach thd.co/update*". In this context, social media represents a useful communication channel. It allows the breached firm to bypass information intermediaries and easily disseminate its intended message (Lee et al. 2015), potentially lowering the cost of the breach (Ponemon Institute 2016) and exposure to litigation risk (Swanson, Kirsch, and Dunigan 2013). Despite these advantages, some studies advise caution in how the value of social media is measured since social media per se cannot generate value without the implementation of an adequate communication strategy and without the allocation of adequate resources (Culnan, McHugh and Zubillaga 2010; Jung et al. 2015). In fact, social media generates an expectation of instant feedback to the public and meeting such expectation may be challenging during a crisis when the number of requests can be overwhelming (Stephens and Malone 2009; Jahng and Hong 2017). Furthermore, due to the virality typical of social media and to the potential high number of information requests, a company may lose control of the information flow, thereby in fact worsening an already serious situation (Blankerspoon et al. 2014; Jung et al. 2015; Lee et al. 2015; Jahng and Hong 2017). Therefore, by increasing corporate organization's exposure to the public

(Conway and Ward 2007), social media may also increase the risk of a potential crisis escalation (Siah Ann Mei, Bansal and Pang 2010). Ultimately, understanding the effect of social media usage around data breaches is critical to providing guidance to firms that suffer data breaches and this study aims to shed light on its potential contrasting impacts.

Our study contributes to the literature on data breaches in two pivotal ways. Firstly, to our knowledge, this is the first study that investigates the impact of communication and disclosure via social media on market reaction to data breach announcements. Previous studies have focused on breach characteristics and firm characteristics. We therefore provide novel and important insights for management into communications strategy around data breaches. Secondly, our study contributes to the ongoing debate on the economic impact of data breach announcements by providing new evidence of a negative price reaction to news of a data breach. Our study also contributes to the literature on the impact of social media for crisis communication by investigating the role of social media in the context of data breach disclosure, and by providing unique evidence of a significant benefit to social media for low visibility firms in particular, which typically struggle in gaining attention in traditional media.

The rest of the paper is organized as follows. The next section presents the theoretical development and the research hypotheses. The following section describes the research design and the data collection. We then present the results of the empirical analyses and the robustness tests, and conclude by discussing the implications of the study and directions for future research.

2. Hypotheses Development

Disseminating information around a breach event quickly is a requirement under compliance with the Security Breach Notification Laws (SBNLs), which have been

enacted in the US since 2003[3]. Academic literature on crisis communication clearly suggests that it is even more important to spread the information to a large audience in order to limit the potential harm and repair any associated reputation damage; to this aim, the use of social media may be particularly beneficial (Seeger 2006). Furthemore, social media allows firms to provide timely updates about the investigation and remediation of the data breach therefore reducing uncertainty and communicating the firm's competence in handling the crisis, hence lessening stakeholders' negative beliefs (Lee et al. 2015).

However, the adoption of social media in the context of a data breach may also have a detrimental effect which may ultimately worsen the stock price reaction. As mentioned above, a company crisis like a data breach is likely generate large volumes of interactions with the public. In this context, keeping control over the information flow may be challenging and at the same time paramount (Lee et al. 2015; Jahng and Hong 2017). If such control is not maintained, the reputation damage may be more pronounced due to the spreading of misinformation around the breach event.

The conflicting arguments outlined so far match with the conclusions of Aral et al. (2013, p. 9), who state that "there is, currently, little understanding with respect to the best ways in which companies should organize and manage social media." In the light of the above, and being the first study investigating the effect of social media usage in the context of data breach disclosure, we are not able to predict what would be the impact of social media and so state our first research hypothesis in null form.

H1: Social media exposure has no impact on the negative stock price reaction to data breach announcements.

[3] Since 2002, when the first SBNL was enacted California, 47 states Forty-seven states, the District of Columbia, Guam, Puerto Rico and the Virgin Islands have enacted their own SNBLs (NCSL 2017). However, SBNLs still significantly differ from each other creating uncertainty in terms of disclosure requirements (Winn 2009; Stevens 2012).

With the first hypothesis establishing whether a relationship exists between general social media exposure and the stock market reaction to data breach announcements, more nuanced effects must next be considered. It is noted that traditional media accommodates high visibility firms, while low visibility firms struggle in reaching a large audience with their company specific news (Miller 2006; Barber and Odean 2008). This may be particularly detrimental when information has to be disclosed quickly, as in the case for data breach events. The risk is that low visibility firms, though they detect the breach quickly, cannot disseminate the event information effectively because they do not command enough attention in traditional media and, therefore, face larger relative damage. Social media levels the playing field somewhat by providing low visibility firms direct access to a potentially wider audience and a greater prospect of market attention than would otherwise be possible (Lee et al. 2015). As such, it provides the affected firm with an opportunity to disclose data breach event information in a more effective manner. Given this important innovation provided by social media, we test whether there is a difference in market reaction between high and low visibility firms. Our contention is that, in contrast to high visibility firms, low visibility firms benefit from having the level of market attention afforded by its social media presence, over and above the case of either no traditional media or very limited traditional media news coverage. For high visibility firms, a social media presence simply adds to an already established level of market attention. We therefore state our second hypothesis as follows.

H2: Social media exposure decreases the negative stock price reaction to data breach announcements for low visibility firms.

While we have thus far focused on general social media communication, we consider next the specific case where social media is used as the outlet to disclose a breach event. Although the disclosure of a data breach is mandatory, under security breach

notification laws, the communication of the event on social media is voluntary. Prior studies show that firms are likely to use voluntary disclosure to share positive news while they are more reluctant to voluntarily disclose bad news (Mayew 2008; Larcker, Larcker, and Tayan 2012). Jung et al. (2015) provide evidence of such a tendency in using social media. Since a data breach is bad news, and with this bad news being broadcast to a wide audience using social media, we expect the negative price reaction to the announcement to be larger if a firm discloses the event through its social media account. Our third hypothesis is therefore stated as follows.

H3: Firm of social media to disclose a data breach increases the negative stock price reaction.

3. Data and Research Methodology

In this study, we adopt an event study methodology (MacKinlay 1997) to investigate whether and how social media exposure affects the stock price reaction to a data breach announcement. The event study methodology is based on the efficient market theory which states that new information in the market will fully reflect in a firm's stock price. Because the market should not be capable of anticipating when firms will make a data breach announcement, it is appropriate to use the event methodology to catch unexpected business events in the stock market (Chai, Kim, and Rao 2011).

4.1 Multivariate Regression Models

The following regression model is used to test H1 and H2:

$$\begin{aligned} CAR_{i,j} = {} & \alpha_0 + \alpha_1 Twitter_{i,j} + \alpha_2 Low\_TMV_{i,j} + \\ & + \alpha_3 Twitter_{i,j} \times Low\_TMV_{i,j} \\ & + \alpha' Controls_{i,j} + \varepsilon_{i,j} \end{aligned} \quad (1)$$

The dependent variable in Equation (1) is the cumulative abnormal return (*CAR*) over a two-day period starting on the announcement day (0; +1) (hereafter, the event period) (Cavusoglu et al. 2004; Gatzlaff and McCullough 2010). We adopt the market model (Fama et al. 1969) to estimate daily abnormal returns (ARs) and then sum up the daily ARs over the event period to obtain the cumulative abnormal returns (CARs), which is proxy for price reaction to the announcement (Campbell et al. 2003; Cavusoglu et al. 2004; Gatzlaff and McCullough 2010). The market model equation is shown in Equation (2).

$$R_{i,t} - RF_t = \alpha_i + \beta_i(RM_t - RF_t) + \varepsilon_{i,t} \quad (2)$$

where $R_{i,t}$ is the stock return for firm i on day t; RF_t is the risk-free interest rate on day t; RM_t is the stock return of market on day t; α_i is Jensen's alpha for firm i,; β_i is the CAPM slope parameter for firm i (i.e., the systematic risk of the return of firm i, relative to the return of the entire market, and often denoted as the beta of the stock); and ε_{it} is the model's error term.

To capture the effect of social media exposure on price reaction, we adopt two indicator variables (i.e. *Twitter* and *Low_TMV*) and the interaction variable between them (*Twitter x Low_TMV*). *Twitter* is equal to 1 when a company had an active Twitter account (Java, Song, Finin, and Tseg 2007) when the breach occurred, and 0 otherwise. *Low_TMV* identifies firms with low visibility on traditional media; as such it is equal to 1 if the average daily number of newspaper articles during the estimation period was below the first tercile threshold, and 0 otherwise. The interaction variable (*Twitter x Low_TMV*) allows us to explore if there is a difference between high and low visibility firms. The regression coefficient of *Twitter* tests H1, while the regression coefficient of the interaction variable tests H2.

A second regression model is used to test whether the use of social media to disclose breach events decreases the negative stock price reaction to data breach announcements (H3). This model is specified as follows:

$$CAR_{i,j} = \alpha_0 + \alpha_1 TweetEvent_{i,j} + \alpha_2 Low\_TMV_{i,j} + \alpha_3 TweetEvent_{i,j} \times Low\_TMV_{i,j} + \alpha' Controls_{i,j} + \varepsilon_{i,j} \quad (3)$$

The dependent variable is still the CARs over the event window, but the variables of interest are different. In order to capture the effect of a data breach announcement trough social media, we create an indicator variable (*TweetEvent*) that denotes whether a firm announced the breach on its Twitter account or not. The interaction variable *TweetEvent x Low_TMV* reveals if the announcement of a breach through a firm's Twitter account generates a differential effect for low visibility companies. The coefficient on this interaction term will provide some complementary insights into H3.

Our models also include four categories of control variables: (a) controls for breach characteristics; (b) controls for traditional media activity; (c) controls for social media activity[4]; and (d) controls for firm characteristics.

We begin with defining the first group of controls. The cost of a data breach, and, therefore, the market reaction to the announcement, depends on the breach type (Campbell et al. 2003; Cavusoglu et al. 2004; Gatzlaff and McCulloug 2010). In this paper, we adopt the classification proposed by the Privacy Rights Clearinghouse[5]. We adopt six different indicator variables to identify breaches due to (a) a payment card fraud (*Card*), (b) an unintended information disclosure (*Disc*), (c) an attack by a hacker (*Hack*), (d) an insider misbehavior (*Insd*), (e) a lost, discarded or stolen portable device

[4] This group of control variables are only included in Equation (3) since we run this regression only on the subsample of firms that had an active Twitter account at the time of the breach.
[5] https://www.privacyrights.org/node/1398.

(*Port*), and (f) an unknown reason (*Unkn*)[6]. In order to avoid collinearity issues, we do not include *Unkn* in the regression model and keep it as baseline. Stock price reaction may also be affected by the number of records breached (Campbell et al. 2003; Cavusoglu et al. 2004; Gatzlaff and McCulloug 2010). However, this information is not always provided at the moment of the announcement. This may be due to the absence of breached records or, more likely, to ongoing investigation; this creates more uncertainty around the overall cost of the breach. We adopt an indicator variable (*RecordsKnown*) that indicates whether a firm disclosed the exact number of breached records when it disclosed the event. We also control for the presence of prior breaches[7] by adding the indicator variable *PriorBreach* since investors might penalize more those firms affected by multiple incidents (Gatzlaff and McCulloug 2010).

Following Lee et al. (2015), we include also a control variable for traditional media and social media activity in our model. *AbnTradMedia* is constructed as presented in Equation (4), where *TradMedia* is the average daily number of newspaper articles during the event period (0,+1) while *NTradMedia* is the average daily number of newspaper articles during a 120-day estimation period ending five days before the event (-125,-6)[8]. This variable provides a measure of the abnormal attention a firm attracts around a data breach announcement.

$$AbnTradMedia_{i,j} = \frac{TradMedia_{i,j} - NTradMedia_{i,}}{NTradMedia_{i,j}} \qquad (4)$$

[6] The Privacy Right Clearinghouse classification includes two more breach categories: physical loss (Phys) if the breach is due to lost, discarded or stolen non-electronic records; and stationary device (Stat) if the breach is due to lost, discarded or stolen stationary electronic devices. Since none of the events in our sample fall into these categories, we do not create any indicator variable for them.

[7] Our approach is similar to Gatzlaff and McCullough (2010), who highlight the need to consider some defined number of years before the sampling period, so as not to bias the results. With our sample running from January 2011 to December 2014, we consider all events reported by Privacy Right Clearinghouse since 2005 in constructing this variable.

[8] While Lee et al. (2015) adopt a 60-day estimation period ending the day before the event, we opt for a longer time window (120 days) in order to have better assess firms' visibility (the shorter the time window, the higher the probability that very low-visibility firms have no media coverage). Furthermore, the estimation period ends five days before the announcement in order to avoid the presence of possible rumors or information leakages. The time between the detection of a data breach and its disclosure may vary considerably since Security Breach Notification Laws do not define precise disclosure timelines and allow delays if a police investigation is underway (Faulkner 2007).

Control variables for social media activity include (a) *AbnTweet* which measures the abnormal Twitter activity of the firm during the event period, and (b) the natural logarithm of the number of followers of the firm's account on the announcement day (*Followers*) which provides a measure of the size of the potential audience. Equation (5) shows how we construct *AbnTweet*. In particular, *Tweet* (*NTweet*) is the average daily number of tweets generated by the Twitter account of firm *i* during the event period (estimation period) of event *j*.

$$AbnTweet_{i,j} = \frac{Tweet_{i,j} - NTweet_{i,j}}{NTweet_{i,j}} \quad (5)$$

We also include control variables for firm characteristics like (a) the size of the affected company measured as the natural logarithm of total assets (*Size*) (Cavusoglu et al. 2004; Gatzlaff and McCullough 2010); (b) its growth opportunities proxied by the market-to-book ratio (*Growth*) (Gatzlaff and McCullough 2010); and (c) whether it operates in an industry with high expectations[9] in terms of cyber security (*HighExp*) (Gatzlaff and McCullough 2010).

Finally, following Lee et al. (2015), we include the interaction variables between *Size* and the other control variables in the model presented in Equation (1) for model specification purposes. Table 1 reports a definition and data source for each variable included in our models.

Insert Table 1 here

[9] According to Gatzlaff and McCullough (2010), banking institutions (SIC codes 6011-6099), insurance firms (SIC codes 6311-6399) and technology companies (SIC codes 7371-7379) fall in this category.

4.2 Sample and Data

We build our sample starting from the list of breaches that occurred from January 2011 to December 2014 as compiled by Privacy Rights Clearinghouse[10]. While the number of Twitter users has increased since 2010 (Lee et al. 2015), the actual number of tweets, which denotes the real activity, only increased dramatically in 2011[11]. For this reason, we adopt 2011 as the starting year of our sample.

The initial event list included 2,257 breaches. Being interested in analyzing the stock price reaction to the announcement, we deleted all events that affected non-publicly traded companies (2,034). We then searched on Lexis-Nexis and Twitter[12] to determine if any information leakage occurred for any of the breaches in our sample within seven days before the official announcement date and, if this was the case, we adjusted the event date to the date of this first newspaper article or tweet. This occurred for 14 events. We also used Lexis-Nexis to check whether any confounding event[13] occurred in a seven-day period before the announcement of a given breach. 57 events were excluded on this basis. In case of multiple events for the same firm, we required the events to be at least 130 days apart from each other. This was necessary to avoid that the event period of an incident falls within the estimation period of a following breach affecting the same firm therefore introducing a bias in the latter. We excluded 47 events that did not meet this condition. In order to ensure a sample of comparable events, we excluded 9 events that were announced during weekends or public holidays. Finally, we

[10] http://www.privacyrights.org/data-breach. This dataset has been adopted in other recently published studies (e.g. Higgs, Pinsker, Smith, and Young 2016; Rosati et al., 2017).
[11] According to Twitter statistics, the number of tweets per day was 35 million in 2010, and 200 million in 2011. See https://blog.twitter.com/2010/measuring-tweets for further details.
[12] We checked for information leakage in the tweets generated from or mentioning the company account.
[13] We consider confounding events all earnings announcements, merger and acquisitions news or rumors, CEO and/or top executive turnover.

excluded 23 events because of missing values. Our final sample therefore consists of 87 events corresponding to 73 individual firms[14].

We searched for the main Twitter accounts or for customer services Twitter accounts on the firms' websites and then used Twitter advanced search to check whether the firms tweeted about the data breach[15]. When a firm had both active main and customer service Twitter accounts, we considered only the customer service accounts as this would more likely be targeted by customers' complaints (Li, Berens, and De Maertelaere 2013). Finally, given that Security Breach Notification Laws were enacted in different years across different states, we checked that all firms in our final sample were subject to mandatory disclosure when the breach occurred. Table 2 summarizes the sampling process. Table 3 provides relative frequencies of events over time, while Table 4 reports the number of events per breach type.

For this study, we retrieved data from three other sources. Daily stock price and market index data was sourced from Thomson Reuters Datastream. We collected the number of newspaper articles using Lexis-Nexis PowerSearch[16]. The Twitter data came from TwitterCounter[17], which provides daily statistics on active Twitter accounts. TwitterCounter statistics include the daily number of tweets generated from a specific account as well as the daily number of followers and followings.

Insert Table 2 here

Insert Table 3 here

Insert Table 4 here

[14] The limited sample size reflects the data availability and the need to apply adequate filters in order to reduce possible noise. Both the sampling criteria and the size of our sample are in line with previous studies on the same topic (Cavusoglu et al. 2004; Gatzlaff and McCulloug 2010; Gordon et al. 2011).

[15] We searched whether any tweet was generated from the official Twitter account containing the following keywords in the event period 'breach OR breached OR breaches OR hacker OR hacked OR attack'. All the tweets retrieved were manually inspected to ensure that they were related to the announcement of the data breach that affected the company that generated the message.

[16] Following Lee et al. (2015), we searched for company name or ticker symbol in the headlines or the lead paragraph of newspaper articles.

[17] http://twittercounter.com/.

4. Results and Discussion

4.1 Univariate Analysis

Table 5 reports the cumulative abnormal returns (CARs) over different time windows. Although the focus of this study is on the most immediate impact of the announcement over the days (0,+1), looking at different time windows is useful to understand how long the effect of the announcement lasts and whether there is a price under-/over-reaction to the announcement. Panel A in Table 5 shows that a data breach has a negative impact over a three-day period starting at the announcement day (0,+2), but the largest price drop occurs over the first two days (0,+1) when breached firms lose on average 1.6 percent of their market value. Results show also that the stock price recovers three days after the announcement, which indicates that a breach causes only a short-term negative effect on stock price. To check for price under- or over-reaction, we also estimated CARs over a seven-day period starting four days after the announcement (from day +4 to day +10). If there was under-(over-)reaction, the stock price would fall (increase) over this period. Panel A shows that CARs from day +4 to day +10 are, on average, positive and statistically different from zero, unveiling a price over-reaction to the announcement. As discussed above, the exact extent (and cost) of a breach is not always clear from its detection and disclosure. This creates uncertainty among investors, which might cause a price drop; such uncertainty might then decrease progressively as firms provide additional information. Panel B, instead, provides a comparison between the two subsamples of firms with and without active Twitter accounts. Results suggest that the two subsamples are not statistically different from one another; the only difference is in the percentage of firms with negative CARs over a three-day event window (0,+2), where 81 (60) percent of firms with (without) Twitter have a negative price change.

Insert Table 5 here

Table 6 shows the descriptive statistics of the variables included in our regression models. We winsorized all continuous variables at 1 and 99 percent to avoid any outlier that could alter the results. Panel A shows that, as reported above, the average price change (CARs) is 1.6 percent. 25 percent of the events in our sample caused a price drop larger than 2.6 percent while another 25 percent caused a moderate increase. The average value of *AbnTradMedia* reveals that traditional media pays a lot of attention to data breaches since the number of newspaper articles concerning the events in our sample increases, on average, by 32 percent over the event period. Looking at social media activity, results show that breached firms increase their Twitter activity, on average, by 10 percent, and that only 9 percent of firms with an active Twitter account decide to disclose the event through their account suggesting opportunistic behavior in social media communication (Jung et al. 2015). Finally, the number of records breached is disclosed for just 35 percent of the events in our sample, while almost 40 percent of the events are preceded by other breaches.

Panel B in Table 6 compares the average values between the two subsamples of firms, those with and without an active Twitter account. Results show that there is a statistically-significant difference only regarding the percentage of low-visibility firms in the two subsamples. Indeed, they represent around 19 (49) percent of observations in the Twitter (Non-Twitter) subsample. This evidence might signal that low-visibility firms did not value the potential benefit of social media communication.

Insert Table 6 here

We also performed a correlation analysis among all the variables included in our regression models. The results (not tabulated) confirm the existence of a positive relationship between company size and traditional media coverage, and show evidence of no strong correlations between different explanatory variables therefore suggesting

that multicollinearity should not affect the results of our regression model (Gujarati and Porter 2003).

4.2 Effect of Social Media Exposure on Stock Price

In order to test the research hypotheses H1-H3, we adopt a pooled ordinary least-squares (OLS) regression with time fixed effects.

Panel A in Table 7 presents OLS coefficients and levels of significance for each the variables of the models presented in Equation (1). This regression involves all 87 events in our sample and the variables of interest are *Twitter* and *Twitter x Low_TMV* since they test H1 and H2 respectively.

The results show that the potential negative effects of social media communication outweigh its potential benefits in the context of data breach announcements. The coefficient of *Twitter* is negative and statistically significant. This indicates that a firm's social media exposure worsens the stock price reaction, on average, by 1.2 percent following a data breach announcement. This leads us to reject H1. This result is the opposite of the one obtained by Lee et al. (2015), who show that the use of interactive social media (i.e. Twitter and Facebook) lowers the negative market reaction to product recall announcements, and provides evidence that different crises need different communication strategies (Utz et al. 2013). The complexity of security breach notifications laws and the greater uncertainty around costs are likely factors for the difference in results.

In contrast to the above, social media exposure for low visibility firms mitigates the negative effect of the breach, on average, by 3.5 percent (see coefficient of *Twitter x Low_TMV*). This result confirms that social media provides low-visibility firms with the opportunity to engage in a more effective communication and supports H2.

Insert Table 7 here

Table 8, instead, presents the OLS coefficients and the levels of significance for the variables of the model presented in Equation (2). This regression is run only on the subsample of 32 events that involved firms with an active Twitter account when the breach occurred. The coefficient of *TweetEvent* tests H3, while the coefficient of *TweetEvent x Low_TMV* provides further evidence for H2.

The results show that the disclosure of a data breach on social media (*TweetEvent*) exacerbates the negative price response to the announcement. In particular, the price drops by 5.2 percent more compared to other companies that have an active Twitter account, but do not disclose the event directly from their account. This leads us to accept H3 and suggests that spreading bad news to a larger audience does not represent a convenient communications strategy in the context of a data breach. However, it seems to be an effective strategy for low visibility firms. The coefficient of the interaction variable (*TweetEvent x Low_TMV*), indeed, shows that the event disclosure through the Twitter account of a low-visibility firm mitigates the negative price response by 4.4 percent. This result provides a further confirmation of H2.

Two other factors are worth attention. Firstly, the abnormal Twitter communication of a breached firm (*AbnTweet*) increases the negative price reaction, on average, by 10 percent. This result is a clear signal that firms tend not to adopt effective communication strategies in their social media usage and/or that they cannot keep (enough) control of the information flow. Secondly, the larger the audience (i.e. followers), the more negative the price response to announcement, as evidenced by the negative coefficient of *Followers*. In particular, the result indicates that stock prices decrease, on average, by 0.53 percent for every 100 followers.

Insert Table 8 here

5. Robustness Test

The dependent variable of both regression models employed in this study is the cumulative abnormal returns (*CAR*). As shown above, we estimate CARs based on the Market Model (Fama et al. 1969). Fama and French (1993) propose an alternative model to estimate CARs. Their model, known as the *Three-factor Model*, includes two factors other than the market index return which are the difference of returns between (a) firms with small and large market capitalization and (b) firms with high and low book-to-market ratio. The *Three-factor Model* equation is shown in Equation (6).

$$R_{i,t} - RF_t = \alpha_i + \beta_i(RM_t - RF_t) + \delta_i SMB_t + \gamma_i HML_t + \varepsilon_{i,t} \quad (6)$$

where $R_{i,t}$ is the stock return for firm i on day t; RF_t is the risk-free rate on day t; RM_t is the return of the market on day t; SMB_t is the difference between the returns on a portfolio of small and large stocks on day t; HML_t is the difference between the return on a portfolio of high and low book-to-market stocks on day t; α_i, β_i, δ_i and γ_i are the model intercept and sensitivity parameters, respectively, for firm i; and $\varepsilon_{i,t}$ is the model's error term.

In order to ensure that the results of our analysis do not depend on the estimation model adopted, we estimate CARs using the *Three-factor Model* and run the regressions using the new CARs as dependent variable. The results (not tabulated) are consistent with ones discussed above, therefore we can conclude that the results of this study are robust to the CAR estimation model specification.

Finally, to ensure that our findings on low visibility firms are not driven by the threshold adopted, we repeat the analysis adopting a quartile-based classification. In this case, low visibility firms are the ones with an average daily number of newspaper articles during the estimation period below the first quartile threshold. The results (not

tabulated) are consistent also in this case, therefore we can conclude that the results of this study are robust to different visibility classification criteria.

6. Conclusion

This paper investigates whether the use of social media affects the price reaction to a data breach announcement. Our empirical analysis suggests that communication via social media (i.e. Twitter) tends to exacerbate the negative impact of data breach announcements on stock price, causing an average additional decrease of 1.2 percent over a two-day event period (0,+1),. Further analyses suggest that the negative effect of social media is even more pronounced when firms (a) disclose the event through their Twitter account (-5.2 percent), (b) increase the communication via social media (i.e. number of tweets) in the event period, and (c) have a larger audience on social media (i.e. followers). However, our results also suggest that the impact of social media is positive for low-visibility firms.

The contribution of this study is threefold. Firstly, the study provides new insights into the cost of data breaches by adding communication via social media as a new significant factor affecting the price reaction to a data breach announcement. In so doing we provide additional evidence on the effectiveness of the use of social media for crisis communication, but we also contribute to the ongoing debate on the net effect that social media generates in crisis communication by providing evidence of a differential impact based on firms' visibility on traditional media. Our study provides evidence and important practical information for firms making communication decisions in crises such as these. Although there is a generalized positive view on the adoption of social media in firms' communications, managers should also be aware of the challenges that it generates, and of the peculiarity of the crisis they are dealing.

Secondly, our paper provides further evidence of a negative price reaction to data breach announcements contributing to the debate about the economic impact of data breaches and showing additional potential outcomes related to the way the information is delivered to the stakeholders.

Thirdly, our study contributes to the research on the impact of company disclosure through social media on stock market by confirming that it significantly affects the stock price and providing evidence of a positive impact on low visibility firms in regard of data breach announcements. By showing that social media usage is likely to either help or hinder a firm in the context of a crisis, these results are likely to be useful for industry as they highlight the need for a contingent crisis communication strategy based on firm visibility and on the type of crisis a firm is facing.

This study is also subject to some limitations. Firstly, our analysis considers only Twitter as a social media platform. Although Twitter is the most accepted platform in the financial community, alternative social media platforms (e.g. Facebook) are available to firms or indeed firms may decide to disclose events through a number of platforms at the same time to reach different stakeholders. The use of alternative platforms, potential interconnections between them, and stakeholders' preferences are not considered in this study, therefore further research on this would be informative. Secondly, our analysis is based on daily statistics about the use of social media. It does not allow us to investigate the content of the messages which might convey more information about the firms' communication strategies. Text mining may provide interesting insights on what type of information breached firms provide. Moreover, an analysis of the communication patterns between breached firms and corresponding stakeholders might reveal how information flows within the network and if breached firms tend to have a proactive (i.e. provide updates on the incidents) or reactive approach (i.e. just reply to customers' or investors' enquiries). Further research in this

field would shed additional light on firms' communication strategy around data breach announcements and bad news disclosure in general.

References

Abbasi, A., S. Sarker, and R.H., Chiang. 2016. Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems* 17(2): 1-32.

Aral, S., C. Dellarocas, and D. Godes. 2013. Introduction to the special issue-social media and business transformation: A framework for research. *Information Systems Research* 24(1): 3-13.

Barber, B.M., and T. Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21(2): 785-818.

Blankespoor, E., S.G. Miller, and H.D. White. 2014. The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *The Accounting Review* 89(1): 79-112.

Campbell, K., L.A. Gordon, M.P. Loeb, and L. Zhou. 2003. The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security* 11(3): 431-448.

Casey, E. 2006. Investigating sophisticated security breaches. *Communications of the ACM* 49(2): 48-55.

Cavusoglu, H., B. Mishra, and S. Raghunathan. 2004. The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *International Journal of Electronic Commerce* 9(1): 70-104.

Chai, S., M. Kim, and H.R. Rao. 2011. Firms' information security investment decisions: Stock market evidence of investors' behavior. *Decision Support Systems* 50(4): 651-661.

Charette, R.N., K.M. Adams, and M.B. White. 1997. Managing risk in software maintenance. *IEEE Software* 14(3): 43-50.

Chen, H., R.H. Chiang, and V.C. Storey. 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly* 36(4): 1165-1188.

Conway, T. and Ward, M. 2007. International crisis potential: the importance of strategic approach to marketing communications. *Journal of Marketing Communications* 13(3): 213-228.

Coombs, W.T. 2007. Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Corporate reputation review* 10(3): 163-176.

Coombs, W.T. 2014. *Ongoing Crisis Communication: Planning, Managing, and Responding: Planning, Managing, and Responding*. Sage Publications.

Culnan, M.J., P.J. McHugh, and J.I. Zubillaga. 2010. How large U.S. companies can use Twitter and other social media to gain business value. *MIS Quarterly Executive* 9(4): 243–259.

Dennis, A., B.H. Wixom, and D. Tegarden. 2015. *Systems analysis and design: An object-oriented approach with UML*. John Wiley & Sons.

Fama, E.F., L. Fisher, M.C. Jensen, and R. Roll. 1969. The adjustment of stock prices to new information. *International economic review* 10(1): 1-21.

Fama, E.F., and K.R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of financial economics* 33(1): 3-56.

Faulkner, B. 2007. Hacking into data breach notification laws. *Florida Law Review* 59(5): 1097-1125.

Federal Trade Commission (FTC). 2009. *Consumer Data Broker ChoicePoint Failed to Protect Consumers' Personal Data, Left Key Electronic Monitoring Tool Turned Off for Four Months*. Available at: https://www.ftc.gov/news-events/press-releases/2009/10/consumer-data-broker-choicepoint-failed-protect-consumers.

Gatzlaff, K.M., and K.A. McCullough. 2010. The effect of data breaches on shareholder wealth. *Risk Management and Insurance Review* 13(1): 61-83.

Goel, S., and H.A. Shawky. 2009. Estimating the market impact of security breach announcements on firm values. *Information & Management* 46(7): 404-410.

Gordon, L.A., M.P. Loeb, and L. Zhou. 2011. The impact of information security breaches: Has there been a downward shift in costs?. *Journal of Computer Security* 19(1): 33-56.

Gujarati, D., and D. Porter. 2003. *Multicollinearity: What happens if the regressors are correlated*, in Basic econometrics 4th Edition: McGraw Hill.

Higgs, J.L., R.E. Pinsker, T.J. Smith, and G.R. Young. 2016. The Relationship between Board-Level Technology Committees and Reported Security Breaches. *Journal of Information Systems* 30(3): 79-98.

Hoffman, D.L., and T.P. Novak. 2012. *Why do people use social media? Empirical Findings and a New Theoretical Framework for Social Media Goal Pursuit*. Working Paper, George Washington University School of Business.

Jahng, M. R., and Hong, S. 2017. How Should You Tweet?: The Effect of Crisis Response Voices, Strategy, and Prior Brand Attitude in Social Media Crisis Communication. *Corporate Reputation Review* 20(2): 147-157.

Jansen B.J., Zhang, M., Sobel, K., and Chowdury, A. 2009. Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology* 60(11): 2169–2188.

Java, A., X. Song, T. Finin, and B. Tseng. 2007. Why we twitter: understanding microblogging usage and communities. In: *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*: 56-65.

Jung, M.J., J.P. Naughton, A. Tahoun, and C. Wang. 2015. *Corporate use of social media*. Working Paper, New York University, Northwestern University.

Kalev, P.S., W.M. Liu, P.K. Pham, and E. Jarnecic. 2004. Public information arrival and volatility of intraday stock returns. *Journal of Banking & Finance* 28(6): 1441-1467.

Kietzmann, J.H., Hermkens, K., McCarthy, I.P., and Silvestre, B.S. 2011 Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons* 54(3): 241-251.

Larcker, D.F., S.M. Larcker, and B. Tayan. 2012. What do corporate directors and senior managers know about social media. In: *The Conference Board Director Notes: The Conference Board Inc.*: 1-15.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., and Kruschwitz, N. 2011. Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 52(2): 21-32.

Lee, L.F., A.P. Hutton, and S. Shu. 2015. The Role of Social Media in the Capital Market: Evidence from Consumer Product Recalls. *Journal of Accounting Research* 53(2): 367-404.

Li, T., Berens, G. and De Maertelaere, M. 2013. Corporate Twitter channels: The impact of engagement and informedness on corporate reputation. *International Journal of Electronic Commerce* 18(2): 97-126.

Liu, X., and Q. Ye. 2016. The different impacts of news-driven and self-initiated search volume on stock prices. *Information & Management* 53(8): 997-1005.

MacKinlay, A.C. 1997. Event studies in economics and finance. *Journal of Economic Literature* 35(1): 13-39.

Mayew, W.J. 2008. Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research* 46(3): 627-659.

Miller, G.S. 2006. The press as a watchdog for accounting fraud. *Journal of Accounting Research* 44(5): 1001-1033.

National Conference of State Legislature (NCSL). 2017. *Security Breach Notification Laws*. Available at: http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx.

Ponemon Institute. 2016. *Cost of Data Breach Study: Global Analysis*. Ponemon Institute. Available at: https://www-01.ibm.com/marketing/iwm/dre/signup?source=mrs-form-1995&S_PKG=ov49542.

Ponemon Institute. 2018. *2018 Cost of a Data Breach Study: Global Overview*. Ponemon Institute. Available at: https://www.pbwt.com/content/uploads/2018/07/2018-Cost-of-Data-Breach-Study.pdf.

Risius, M., and Beck, R. 2015. Effectiveness of corporate social media activities in increasing relational outcomes. *Information & Management* 52(7): 824-839.

Romanosky, S., D. Hoffman, and A. Acquisti. 2014. Empirical analysis of data breach litigation. *Journal of Empirical Legal Studies* 11(1): 74-104.

Rosati, P., P. Deeney, F. Gogolin, M. Cummins, L. van der Werff, and T. Lynn. 2017. The effect of data breach announcements beyond the stock price: Empirical evidence on market activity. *International Review of Financial Analysis* 49: 146-154.

Schultz, F., S. Utz, and A. Göritz. 2011. Is the medium the message? Perceptions of and reactions to crisis communication via twitter, blogs and traditional media. *Public relations review* 37(1): 20-27.

Seeger, M.W. 2006. Best practices in crisis communication: An expert panel process. *Journal of Applied Communication Research* 34(3): 232-244.

Securities and Exchanges Commission (SEC). 2013. *SEC Says Social Media OK for Company Announcements if Investors Are Alerted*. Available at: https://www.sec.gov/News/PressRelease/Detail/PressRelease/1365171513574.

Sen, R., and S. Borle. 2015. Estimating the Contextual Risk of Data Breach: An Empirical Approach. *Journal of Management Information Systems* 32(2): 314-341.

Siah Ann Mei, J., Bansal, N., and Pang, A. 2010. New media: a new medium in escalating crises?. *Corporate Communications: An International Journal* 15.2: 143-155.

Srinidhi, B., J. Yan, and G.K. Tayi. 2015. Allocation of resources to cyber-security: The effect of misalignment of interest between managers and investors. *Decision Support Systems* 75: 49-62.

Stephens, K.K., and P.C. Malone. 2009. 'If the organizations won't give us information: The use of multiple new media for crisis technical translation and dialogue. *Journal of Public Relations Research* 21(2): 229–239.

Stevens, G. 2012. *Data Security Breach Notification Laws*. Congressional Research Service. Available at: https://pdfs.semanticscholar.org/8f37/2875c6cdc54a0c40b65180a6b117bc228d61.pdf.

Stieglitz, S., and L. Dang-Xuan. 2013. Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems* 29(4): 217-248.

Swanson, K., Kirsch, T.L., and Dunigan, R.M. 2013. *Data breaches in a whistleblower's world: What you should know, why you should know it*. PricewaterhouseCoopers LLP. Available at: https://www.pwc.com/us/en/forensic-services/publications/assets/data-breaches-whistleblowers-world.pdf.

Utz, S., F. Schultz, and S. Glocka. 2013. Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review* 39(1): 40-46.

Verizon. 2017. *Data Breach Digest: Perspective is Reality*. Available at: http://www.verizonenterprise.com/verizon-insights-lab/data-breach-digest/2017/.

Winn, J. K. 2009. Are "better" security breach notification laws possible?. *Berkeley Technology Law Journal* 24: 1133-1166.

Table 1. Variable Description and Data Sources

Variable	Description	Source
CARs(0,1)	Cumulative abnormal returns calculated as the sum of the differences between the actual daily stock returns and the expected stock returns (estimated through the Market Model) during the event period.	DataStream Professional
Twitter	The variable equals 1 if the firm has an active Twitter account when the data breach occurred.	Individual Twitter accounts
Low_TMV	The variable equals 1 if the firm has an average daily number of newspaper articles during the estimation period below the first tercile threshold.	Lexis-Nexis
Twitter x Low_TMV	Interaction variable between *Twitter* and *Low_TMV*.	
TweetEvent	The variable equals 1 if the firm posts a tweet announcing the data breach during the event period.	Individual Twitter accounts
TweetEvent x Low_TMV	Interaction variable between *TweetEvent* and *Low_TMV*.	
AbnTweet	Abnormal Twitter activity measured as the difference between the average daily number of tweets during the event and the estimation period, scaled by the average daily number of tweets during the estimation period.	TwitterCounter
Followers	Natural logarithm of the number of followers of the Twitter account of the firm on the announcement day.	TwitterCounter
Card	The variable equals 1 if the breach is due to payment card fraud.	Privacy Rights Clearinghouse
Disc	The variable equals 1 if the breach is due to unintended information disclosure.	Privacy Rights Clearinghouse
Hack	The variable equals 1 if the breach is due to a hacker attack.	Privacy Rights Clearinghouse
Insd	The variable equals 1 if the breach is due to insider misbehavior.	Privacy Rights Clearinghouse
Port	The variable equals 1 if the breach is due to a lost, discarded or stolen portable device.	Privacy Rights Clearinghouse
Unkn	The variable equals 1 if the cause of the breach is unknown.	Privacy Rights Clearinghouse
RecordsKnown	The variable equals 1 if the number of breached records is disclosed at the first announcement.	Privacy Rights Clearinghouse
PriorBreach	The variable equals 1 if the firm had other breach(es) before the event.	Privacy Rights Clearinghouse

AbnTradMedia	Abnormal number of newspaper articles calculated as the difference between the average daily number of newspaper article during the event and the estimation period, scaled by the average daily number of newspaper articles during the estimation period.	Lexis-Nexis
Size	Natural logarithm of firm's total assets at the end of the fiscal year before the event.	DataStream Professional
Growth	The ratio between the book and the market value of firm's equity at the end of the fiscal year before the event.	DataStream Professional
HighExp	The variable equals 1 if the firm operates in the banking, insurance of tech industries (SIC codes 6011-6099, 6311-6399, 7371-7379).	DataStream Professional

Table 2. Sample Definition

	Number of events		
Filters	Full Sample	Twitter	No-Twitter
Events reported by Privacy Rights Clearinghouse (2011-2014)	2,257		
Non-publicly traded firms	(2,034)		
Events with possible confounding announcements	(57)		
Events overlapping	(47)		
Announcement during weekends or public holidays	(9)		
Missing data	(23)		
Final Sample	87	32	55
Number of firms	73	29	44

This table describes the sampling process. The event is defined as the data breach announcement reported by Privacy Rights Clearinghouse or the first newspaper article reporting the event. The final sample includes 87 data breaches affecting 73 individual firms. Columns 'Twitter' and 'No-Twitter' report the number of data breaches and firms with or without an active Twitter handle respectively.

Table 3. Events Distribution over Time

Year	Full Sample		Twitter		No-Twitter	
	No. of Events	%	No. of Events	%	No. of Events	%
2011	26	29.89%	8	9.20%	18	20.69%
2012	19	21.84%	9	10.34%	10	11.49%
2013	30	34.48%	8	9.20%	22	25.29%
2014	12	13.79%	7	8.05%	5	5.75%
TOT	87	100.00%	32	36.78%	55	63.22%

This table provides the distribution of the events over time for the full sample as well as for the subsamples of firms with (Twitter) and without (No-Twitter) an active handle.

Table 4. Events Distribution per Type of Data Breach

Type of Breach	Full Sample		Twitter		No-Twitter	
	No. of Events	%	No. of Events	%	No. of Events	%
Payment card Fraud	6	6.90%	2	2.30%	4	4.60%
Disclosure	14	16.09%	2	2.30%	12	13.79%
Hacker	26	29.89%	15	17.24%	11	12.64%
Insider	17	19.54%	7	8.05%	10	11.49%
Portable device	5	5.75%	3	3.45%	2	2.30%
Unknown	8	9.20%	3	3.45%	5	5.75%
TOT	87	100.00%	32	36.78%	55	63.22%

This table provides distribution of the events per type of data breach for the full sample as well as for the subsamples of firms with (Twitter) and without (No-Twitter) an active handle. Breach categories correspond to the ones proposed by Privacy Rights Clearinghouse.

Table 5. Cumulative Abnormal Returns Analysis

Panel A: Cumulative Abnormal Return Analysis for the Full Sample over Different Time-Windows

Time Windows	Mean	p-Value (H0=0)	Median	p-Value (H0=0)	Standard Deviation	Percentage of negative CARs	p-Value (H0=50%)
(0,1)	-0.016	0.000 ***	-0.014	0.000 ***	0.027	72.20%	0.000 ***
(0,2)	-0.008	0.029 **	-0.008	0.002 ***	0.036	66.00%	0.001 ***
(0,3)	0.010	0.025 **	0.009	0.025 **	0.041	38.10%	0.019 **
(4,10)	0.037	0.000 ***	0.031	0.000 ***	0.080	29.90%	0.000 ***

This panel provides CARs' mean value, median value, standard deviation and percentage of negative CARs together with p-Values associated with the t-Tests of their significance. *, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent or lower levels, respectively.

Panel B: Cumulative Abnormal Return Analysis for the Twitter and No-Twitter Subsamples over Different Time-Windows

	Mean			Percentage of negative CARs		
Time Windows	Twitter	No-Twitter	p-Value	Twitter	No-Twitter	p-Value
(0,1)	-0.018	-0.014	0.512	78.13%	70.91%	0.468
(0,2)	-0.016	-0.007	0.254	81.25%	60.00%	0.041 **
(0,3)	0.001	0.011	0.294	50.00%	32.73%	0.114
(4,10)	0.018	0.043	0.156	31.25%	27.27%	0.697

This panel provides CARs' mean value, percentage of negative CARs for Twitter and No-Twitter Subsamples. The p-Values are the level of significance of the t-Tests under the null hypothesis of equal mean in the two subsamples. *, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent or lower levels, respectively.

Table 6. Descriptive Statistics

Panel A: Sample Descriptives

Variable	n	Mean	Std. Dev.	P25	Median	P75
CAR(0,1)	87	-0.016	0.027	-0.026	-0.014	0.001
Size	87	9.716	1.946	8.331	10.008	11.344
Growth	87	0.282	7.110	0.193	0.405	0.961
AbnTradMedia	87	0.326	0.733	0.036	0.119	0.384
AbnTweet	32	0.101	0.061	0.046	0.103	0.138
Followers	32	11.703	1.882	10.955	11.685	12.285
TweetEvent	32	0.094	0.296			
Low_TMV	87	0.333	0.488			
RecordsKnown	87	0.356	0.482			
PriorBreach	87	0.391	0.491			
HighExp	87	0.264	0.444			

This panel provides descriptive statistics for the main variables included in the regression analysis. Descriptive statistics include mean (Mean), standard deviation (Std. Dev.), first quartile (P25), median (Median), and third quartile (P75). All the variables are presented in Table 1.

Panel B: Sample Descriptives for Twitter and No-Twitter Subsamples

Variable	Twitter	No-Twitter	p-Value	
CAR(0,1)	-0.018	-0.014	0.512	
Size	9.978	9.564	0.342	
Growth	-0.860	0.947	0.256	
AbnTradMedia	0.362	0.305	0.726	
Low_TMV	0.188	0.491	0.005	***
RecordsKnown	0.375	0.345	0.784	
PriorBreach	0.438	0.364	0.502	
HighExp	0.188	0.309	0.220	
n	32	55		

This panel provides the mean values of the main variables included in the regression analysis for the Twitter and No-Twitter subsamples. P-values denote the level of significance of the t-Tests under the null hypothesis of equal means in the two subsamples. *, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent, respectively. All the variables are presented in Table 1.

Table 7. Regression Results: General Use of Social Media

Dependent Variable: CARs(0,1)						
Variable	E[sign]	Coefficient	Std. Err.	t-Stat	p-Value	
Intercept		-0.0172	0.048	-0.36	0.721	
Twitter	+/-	-0.0119	0.007	-1.69	0.096	*
Low_TMV		-0.0103	0.009	-1.15	0.256	
Twitter x Low_TMV	+	0.0350	0.011	3.17	0.002	***
Card		-0.2342	0.185	-1.27	0.210	
Disc		0.0771	0.053	1.45	0.154	
Hack		0.0350	0.045	0.79	0.435	
Insd		0.0602	0.057	1.05	0.298	
Port		0.0662	0.051	1.30	0.197	
RecordsKnown		-0.0211	0.038	-0.55	0.582	
PriorBreach		-0.0145	0.043	-0.34	0.739	
AbnTradMedia		-0.0003	0.000	-1.11	0.271	
Size		0.0015	0.005	0.27	0.791	
Growth		0.0051	0.003	1.77	0.081	*
HighExp		-0.2574	0.079	-3.26	0.002	***
Card X Size		0.0187	0.016	1.17	0.248	
Disc X Size		-0.0064	0.006	-1.07	0.288	
Hack X Size		-0.0038	0.005	-0.74	0.463	
Insd X Size		-0.0046	0.006	-0.75	0.455	
Port X Size		-0.0064	0.006	-1.07	0.290	
RecordsKnown X Size		0.0021	0.004	0.58	0.565	
PriorBreach X Size		0.0020	0.004	0.48	0.636	
AbnTradMedia X Size		0.0000	0.000	1.29	0.203	
Growth X Size		-0.0010	0.001	-1.86	0.067	*
HighExp X Size		0.0245	0.008	3.08	0.003	***
Year Fixed Effects			Yes			
R-squared			0.49			
F-Stat			6.35			
p-Value			0.000			
N			87			

This table provides OLS coefficients, heteroskedasticity-consistent standard errors, t-statistics and p-values for the variables included in the regression model presented in Equation (1). *, **, *** indicate significance at the 10 percent, 5 percent, and 1 percent, respectively. All the variables are presented in Table 1.

Table 8. Regression Results: Specific Use of Social Media

Dependent Variable: CARs(0,1)						
Variable	E[sign]	Coefficient	Std. Err.	t-Stat	p-Value	
Intercept		0.0680	0.023	2.94	0.011	**
TweetEvent	-	-0.0520	0.008	-6.20	0.000	***
TweetEvent x Low_TMV	+	0.0443	0.013	3.36	0.005	***
AbnTweet		-0.1318	0.056	-2.35	0.035	**
Followers		-0.0061	0.002	-3.43	0.004	***
Card		-0.0423	0.008	-4.99	0.000	***
Disc		0.0221	0.012	1.85	0.088	*
Hack		0.0145	0.007	1.97	0.070	*
Insd		0.0179	0.009	1.96	0.072	*
Port		-0.0126	0.015	-0.86	0.405	
RecordsKnown		0.0063	0.007	0.87	0.398	
PriorBreach		-0.0073	0.008	-0.92	0.376	
AbnTradMedia		0.0002	0.000	4.60	0.000	***
Size		-0.0010	0.002	-0.59	0.564	
Growth		0.0003	0.000	2.70	0.018	**
HighExp		0.0251	0.015	1.71	0.112	
Year Fixed Effects			Yes			
R-squared			0.82			
F-Stat			26.41			
p-Value			0.000			
N			32			

This table provides OLS coefficients, heteroskedasticity-consistent standard errors, t-statistics and p-values for the variables included in the regression model presented in Equation (2). *, **, *** indicate significance at the 10 percent, 5 percent, and 1 percent, respectively. All the variables are presented in Table 1.